J. Clin. Chem. Clin. Biochem.
Vol. 17, 1979, pp. 283–297

Reference Materials and Reference Methods in Clinical Chemistry[1])

By *D. Stamm*

Department of Clinical Chemistry, Max-Planck-Institut für Psychiatrie, Munich

(Received January 8, 1979)

Summary: In clinical chemistry today there are often several different methods and instruments in use for the analysis of a particular constituent. When these different methods and instruments are used for careful analysis of the same control specimen, as in assigned value determinations, the assigned values obtained may be significantly different, and these differences may be of clinical importance. In order to determine the reasons for such differences, reference systems or reference points are needed. The reference systems for calibration (calibration materials, standard reference materials) and control (control materials), where the matrices of these different materials must have quite different characteristics, are referred to jointly as reference materials. In order to compare methods, reference methods are needed with known, high reliability. Because of the amount of specimen material needed, the time needed for analysis and the facilities required, reference methods are not suitable for routine analysis. Ideally, standard reference materials are developed first, followed by a definitive method, which is then used to evaluate a candidate reference method. This is currently the case for only a small number of the constituents analyzed in the clinical chemistry laboratory. In this paper the characteristics of the different kinds of reference materials and of reference methods are described. The differences in terminological usage found in the literature are discussed.

Limitations of knowledge and technique may necessitate certain compromises, with respect to the ideal characteristics of reference materials and reference methods. Possible compromises are discussed, and the various sources of error associated with them are pointed out.

However, these compromises can also lead to improvement in the reliability and comparability of analytical results from different laboratories. As an example of such improvement, the results are presented for assigned value determinations by highly qualified reference laboratories on control specimens for interlaboratory surveys; the decision limits are included and the results are compared with those of the survey participants.

Thus various ways are indicated for the use of reference materials and reference methods to improve the reliability and comparability of analytical results, suited to the current state of the art.

Referenzmaterialien und Referenzmethoden in der Klinischen Chemie

Zusammenfassung: Derzeit werden in der Klinischen Chemie für die Untersuchung desselben Bestandteiles häufig eine größere Zahl verschiedener Methoden und Geräte benutzt, die bei sorgfältiger Analyse derselben Kontrollprobe zu signifikant verschiedenen Sollwerten führen, deren Unterschiede diagnostisch bedeutsam sind. Für die Aufklärung dieser Unterschiede und ihrer Ursachen werden Bezugssysteme benötigt. Die Bezugssysteme für die Kalibration (calibration materials, standard reference materials) und für die Kontrolle (control materials), die sehr unterschiedliche Matrix-Eigenschaften haben sollen, werden unter dem Begriff Reference Materials zusammengefaßt. Für den Methodenvergleich werden Referenzmethoden mit genau bekannter hoher Zuverlässigkeit benötigt; sie sind infolge der benötigten Probemenge, Analysendauer und Einrichtungen nicht für Routineuntersuchungen geeignet. Im Idealfall werden nacheinander Standard Reference Materials und eine Definitive Method entwickelt, an der eine Reference Method dann geprüft wird. Dieser Idealfall ist derzeit nur bei einem kleinen Teil der in der Klinischen Chemie untersuchten Bestandteile gegeben. In der vorliegenden Arbeit werden die Charakteristika der verschiedenen Referenzmaterialien und der Referenzmethoden beschrieben. Der unterschiedliche Inhalt und Gebrauch der Begriffe in der Literatur wird gegenübergestellt.

[1]) Based on a talk given at a plenary session of the conference "Biochemische Analytik 78" in Munich, April 18–21, 1978, in honour of Professor *Adolf Butenandt* on the occasion of his 75th birthday.

0340-076X/79/0283-0297$2.00

Es wird gezeigt, welche Kompromisse gegenüber den Idealforderungen aufgrund unserer derzeit begrenzten Kenntnisse und technischen Möglichkeiten eingegangen werden. Auf die mit diesen Kompromissen verbundenen Irrtumsmöglichkeiten wird hingewiesen.

Diese Kompromisse können aber auch eine Verbesserung der Zuverlässigkeit und der Vergleichbarkeit der Analysenergebnisse aus verschiedenen Laboratorien bewirken.

Als Beispiele dafür werden die Ergebnisse der Sollwert-Ermittlung in Kontrollproben in qualifizierten Referenzlaboratorien und die Festlegung der Bewertungsgrenzen für Ringversuchsproben mit Routinemethoden mitgeteilt und den Ergebnissen der Teilnehmer gegenübergestellt.

Damit werden, angepaßt an den jeweiligen Stand der Kenntnisse und technischen Möglichkeiten, verschiedene Wege für den Einsatz von Referenzmaterialien und Referenzmethoden zur Erzielung zuverlässiger und vergleichbarer Analysenergebnisse aufgezeigt.

Introduction

All clinical chemical findings, whether used for research purposes or as part of patient care, should be based on reliable and comparable analytical results. The required reliability and comparability of analytical results can be assured only if suitable reference systems are available for calibration and control of the analyses and for comparison of analytical methods.

The reference systems or reference points for calibration and control are referred to as reference materials (1), and those for the comparison of methods as reference methods (2, 3). Unfortunately these terms mean different things to different authors (1, 8). A discussion of differences in usage is urgently needed, and this is one of the tasks of this paper. Interwoven into this discussion are suggestions by the author on standardization of terminology.

No new formal model is presented, however. Rather, the discussion is based on the actual situation in clinical chemistry today. It includes an examination of the characteristics of optimal reference materials and reference methods, a discussion of the compromises necessary in light of our current knowledge and technical possibilities, and the extent to which these compromises still permit reliable analytical results.

Following a discussion of the particular problems associated with clinical chemical analysis there are separate sections on reference materials and reference methods, including their function, the resulting requirements and the different kinds of reference materials and methods. Because of the many ways in which reference materials and reference methods influence each other, numerous references to preceding and following sections are necessary. Practical experience with reference materials and reference methods and the improvement in reliability and comparability which has been made on the basis of this experience are discussed in the final section.

On November 16 and 17, 1977, a conference was held in Atlanta, Georgia, USA on "A National Understanding for the Development of Reference Materials and Reference Methods for Clinical Chemistry". The participants were experts from the relevant federal government agencies, professional associations and health industries in the United States. All participants received position papers (e. g. l. c. (4–7)) by mail prior to the conference, part of the excellent conference planning. The results of the discussions at this conference have been taken into consideration in the present paper.

1. Problems in Clinical Chemical Analysis

Most of the analytical methods used in clinical chemistry involve measurements of a relative nature, i. e., the analytical result from a patient specimen is determined by comparing the reading made on this patient specimen with the reading made on a *standard* of known concentration. Only highly purified, defined standards (8) should be used as the basis of such relative measurements.

In clinical chemistry it often happens that the specimen size and the time available for analysis are so limited that it is not possible to carry out the separation operations necessary for a completely specific determination. As a result, only a very few completely specific methods are available for routine determinations; in most cases the analytical results from patient specimens include the influence of nonspecific components.

An analytical assessment of analytical results is therefore possible only if all of the reliability criteria (2, 3, 9) (tab. 1-1) for the method used are evaluated during the development of the method, described when the method is published and reevaluated when an individual laboratory begins to use the method. Beyond this, precision, accuracy and specificity along with its possible influence on accuracy require regular monitoring through an effective system of quality control (13–16); suitable *control specimens* are needed for this purpose. They must be as similar to the patient specimens as possible (8), i. e. they must include the same nonspecific components, if monitoring is to be effective. The analytical results obtained from optimal standards, optimal control specimens and patient specimens can be expected to be affected by different components (tab. 1-2).

In order to evaluate or compare analytical methods and to estimate the "true value" of constituents in control specimens, particularly reliable analytical

Tab. 1-1. Reliability criteria.

Criterion (Reference)	Definition
Precision (2, 6)	Agreement between replicate measurements. It has no numerical value.
Imprecision (2, 6)	Standard deviation or coefficient of variation of the results in a set of replicate measurements. The mean and number of replicates must be stated, and the design (within-day, between-day, between-laboratory) described.
Accuracy (2, 6)	Agreement between the best estimate of a quantity and its "true value". It has no numerical value.
Inaccuracy (2, 6)	Numerical difference between the mean of a set of replicate measurements and the "true value". This difference (positive or negative) may be expressed in the units in which the quantity is measured, or a percentage of the "true value".
Specificity (2, 9–11, 13)	The ability of an analytical method to determine solely the component(s) it purports to measure.
Detection limit (12, 13)	Analytical result which is clearly detectable and different from the background noise defined as three standard deviations of the appropriate blank value.

methods are required. Among other things, they should allow separation of all nonspecific components. Such methods, usually too complicated to be used in routine analysis, are called *reference methods*.

2. Reference Materials

According to a tentative definition of the International Organization for Standardization (ISO), a reference material is "a material or substance one or more properties of which are sufficiently well established to be used for the calibration of an apparatus or for the verification of a measurement method" (17).

According to this definition, the term reference material includes both standards and control specimens.

Matrix

Most reference materials contain one or more components in addition to the analyte. In the case of a primary standard solution it is the pure diluent, whereas with control specimens there are many components, both major and minor. The sum of the major and minor components and their structures, in which the analyte is embedded, i. e. the environment surrounding the analyte, is referred to as the *matrix* (4).

It is useful here to differentiate among various kinds of matrix (4) (tab. 2-1).

Basic requirements (8)

1. Standards and control specimens must be used completely independently of one another.
2. Standards should be as highly purified and well defined as the state of the art allows. Ideally, solutions would be prepared using well-defined solvents only.
3. The matrix of the control specimens should be as much like the matrix of the specimens to be examined as possible so that procedural deficiencies and interference factors due to the matrix of the specimens can be monitored.

2.1 Standards

A prerequisite for a reliable measurement is knowledge of the exact amount of the analyte in the standard or standard solution. It is customary to distinguish between two types of standards (8, 13) and solutions made from them (tab. 2-2).

In clinical chemistry it is difficult to obtain primary standards which meet the specifications of the International Union of Pure and Applied Chemistry (IUPAC) (19). Because of this the National Bureau of Standards (NBS) in the United States has recently begun supplying primary standards (1, 20) with official certificates which are suitable for clinical chemical analysis (tab. 2-3) or for monitoring spectrometers and thermometers.

Tab. 1-2. Components of analytical results.

Solutions (Index)	Reading (E)	Components				
		Best estimate of "true value" (T)	Deficiencies in procedure (D)	Nonspecific component (N)	Interference component (I)	Result, e. g. concentration
Primary standard solution (S)	E_S	c_{TS}	$\pm c_{DS}$			$= c_S$
Control specimen (C) (Standard specimen)	E_C	c_{TC}	$\pm c_{DC}$	$\pm c_{NC}$		$= c_C$
Patient specimen (P)	E_P	c_{TP}	$\pm c_{DP}$	$\pm c_{NP}$	$\pm c_{IP}$	$= c_P$

Tab. 2-1. Types of matrix in standards and control specimens.

1. Synthetic matrix
   - 1.1 Pure substance(s) in pure solvent(s)
   - 1.2 Pure substance(s) in pure albumin solution
2. Semisynthetic matrix
   - 2.1 Isolated enzyme in pure albumin solution
   - 2.2 Serum preparation spiked with pure substance(s)
   - 2.3 Combination of serum fractions
   - 2.4 Native serum spiked with pure substance(s)
3. Biological matrix
   - 3.1 Native serum spiked with serum fractions
   - 3.2 Native serum pool

Tab. 2-2. Types of standards and standard solutions.

Standard or Standard Material is a substance of known purity which is used as the basis for a comparison of measurements in the determination of this substance or another substance in a specimen.

Primary Standard:	Primary Standard Solution:
Mass can be determined exactly by weighing the pure substance.	Standard solution of known concentration. Produced by weighing the primary standard and dissolving it in a suitable pure or defined solvent.
Secondary Standard:	**Secondary Standard Solution:**
Mass can be determined only indirectly, by chemical analysis.	Standard solution produced by dissolving a secondary standard in a suitable solvent. Determination of concentration by chemical analysis.
	Standard Specimen:
	A secondary standard which contains the same major and minor components as patient specimens. Determination of concentration by chemical analysis.

A reading taken on an optimal standard has no matrix component. If in addition to the defined substances a standard also contains other known or unknown substances or if its matrix is similar to the specimens examined, then it is called a *standard specimen.*

If such standard specimens are used for the calibration of methods which are not completely specific, there is a good deal of uncertainty in the measurements: the known and unknown nonspecific components contribute to the reading taken on the standard, and thus the results found for the patient specimens are determined in part with reference to undefined or unknown components of the matrix. Experience has shown that the kind and quantity of nonspecific components in standard specimens can vary markedly from lot to lot.

2.2 Control specimens

Every quantitative analysis in clinical chemistry must be subjected to the best quality control that is possible on the basis of our present knowledge. A simple but effective basic program was designed for this purpose and is laid down in the Guidelines of the Medical Society of West Germany (Bundesärztekammer) (14, 15) (tab. 2-4).

Control specimens are required for all three parts of this program. The matrix of these control specimens plays a major role in the effectiveness of the monitoring procedures in which the control specimens are used. The closer the matrix of the control specimen is to that of the patient specimens the more effective the control procedures are (8).

The different kinds of control specimens currently in use are shown in table 2-5.

The criteria that control specimens should meet, and the testing necessary to insure specimens of high quality

Tab. 2-3. Some primary standards currently available from NBS with Standard Reference Material (SRM) order number.

SRM	Type	Purity (%)
911 a	Cholesterol	99.4
912	Urea	99.7
913	Uric Acid	99.7
914	Creatinine	99.8
915	Calcium Carbonate	99.9
916	Bilirubin	99.0
917	*D*-Glucose	99.9
918	Potassium Chloride	99.9
919	Sodium Chloride	99.9
920	*D*-Mannitol	99.8
921	Cortisol	98.9
924	Lithium Carbonate	100.0
925	4-Hydroxy-3-methoxymandelic acid (VMA)	99.4
928	Lead Nitrate	100.0
929	Magnesium Gluconate	*

* in preparation

Tab. 2-4. Quality control according to the Guidelines of the Medical Society of West Germany.

1. *Internal quality control*	
a) Control of precision	in every run of analyses, even if the run consists of only one patient specimen; with precision control specimen
b) Control of accuracy	in at least every 4 th run; with accuracy control specimen
2. *External quality control*	
Interlaboratory surveys	must be run at least three times a year for each constituent.
	specimens shipped by the chief investigator
	evaluation on the basis of analytical results from independent reference laboratories

Tab. 2-5. Types of control specimens.

Types of control specimens (matrix)
1. Pure solutions (synthetic matrix)
1.1 pure substance(s) in pure solvent(s)
1.2 pure substance(s) in pure albumin solution
2. Liquid control specimens (semisynthetic matrix)
2.1 isolated enzyme in pure albumin solution
2.2 serum preparations spiked with pure substance(s)
2.3 combinations of serum fractions
3. Lyophilized control specimens reconstituted prior to use, limited stability after reconstitution
3.1 native serum spiked with pure substance(s) (semisynthetic matrix)
3.2 native serum spiked with serum fractions (biological matrix)
4. Frozen control specimens (biological matrix)
limited stability after thawing
4.1 native urine and serum specimens
4.2 spiked native urine or serum specimens

are described in detail elsewhere (8, 21, 22). Therefore only the most important points are mentioned here:

1. The concentration of the constituent being monitored must be in a range where precise measurement is possible and which is also relevant to the intended use of the analytical results.
2. Intervial variability in the concentration of the control specimen should be negligible compared to the precision of the analysis.
3. The concentration of the constituents to be monitored should be stable for a long time under proper storage conditions in an unopened container.
4. Detailed information should be available on the stability of the concentration of relevant constituents after the specimen has been thawed or reconstituted or simply opened.

Liquid control specimens with a semisynthetic matrix are the easiest to handle. However, this type of matrix often does not simulate the natural matrix well enough to assure adequate monitoring of all of the important interference factors resulting from the natural matrix (8).

A number of constituents and the natural matrix are stable for a sufficient length of time in lyophilized form only. The *lyophilized material* must be dissolved (reconstituted) with great care prior to use. The concentration of certain constituents changes more rapidly in reconstituted specimens than in native specimens. Special protection against outside influences (e. g. against light for bilirubin and creatine kinase) may be necessary. The situation is analogous for frozen specimens after thawing.

Control of precision

Control specimens for precision control should have a particularly long shelf life because they are used, among other things, to monitor the long-term stability of analytical systems. Furthermore, whenever the lot of the precision control specimen is changed a preperiod of 20 working days is required to set up a new control chart (14, 15).

The concentration of the constituents to be monitored should be close to the most frequent decision limits for those constituents (15). If the decision limit is near the detection limit, then a concentration should be selected which is close to this decision limit but still in a range which can be measured precisely (8); examples are bilirubin and creatinine in serum and certain enzyme activities in serum (aspartate and alanine aminotransferase).

Control of accuracy and interlaboratory surveys

For both of these kinds of monitoring control specimens are needed in which the concentration of the analytes is known as accurately as possible; the value should be a good estimate of the "true value". For according to the definition of the Expert Panel on Nomenclature and Principles of Quality Control in Clinical Chemistry of the International Federation of Clinical Chemistry (IFCC EP-NPQC) (2), inaccuracy is the deviation from the best estimate of the "true value". This good estimate of the "true value" is obtained by means of replicate determinations with what is called a reference method. For a definition of the term reference method and a discussion of such methods, see Section 3. For analytical methods which are specific this kind of accuracy control is certainly an effective way to monitor for systematic errors.

But since, as has already been mentioned, analytical results obtained with routine methods may contain nonspecific components of various sizes (table 1-2), marked differences are seen between the best estimate of the "true value" and the location parameter (e. g. mean, median) of analytical results found with a routine method which is under control. This location parameter of analytical results found with a routine method is called the *assigned value*. The differences between a good estimate of the "true value" and the assigned values obtained with various routine methods are shown in table 2-6.

Thus the best estimate of the "true value" is not very well suited for monitoring nonspecific routine methods for systematic errors. The assigned value would appear to be a better point of reference for nonspecific routine methods or those with procedural deficiencies.

The method for assigned value determination successfully used in the Federal Republic of Germany over the last several years (23) is described in detail in Section 4. In interlaboratory surveys decision limits are needed in addi-

Tab. 2-6. Comparison of assigned values using different methods.

Constituent (unit) Method	Assigned value (95% confidence limits)	
Creatinine (mg/100 ml)	Reg. No. 16900 lyophilized	Reg. No. 17100 lyophilized
Jaffé reaction		
– after precipitation of the protein and specific adsorption on Fuller's earth	1.12(1.01–1.23)	1.65(1.51–1.79)
– kinetic	0.95(0.85–1.05)	1.5 (1.2 –1.7)
– Autoanalyzer	1.2 (1.1 –1.3)	1.7 (1.5 –1.9)
– after precipitation of the protein with trichloroacetic acid	1.39(1.27–1.51)	1.94(1.77–2.11)
Total protein (g/l)	Reg. No. 16900 lyophilized	Reg. No. 16600 liquid
Kjeldahl method	51.5(50.5–52.5)	–
Biuret reaction		
– without sample blank	56.0(52.0–60.0)	55.0(51.0–59.0)
– with sample blank	52.0(48.0–56.0)	55.0(51.0–59.0)

tion to the assigned value. The determination of these limits is also discussed in Section 4.

2.3 Kinds of reference materials

The NBS has established a hierarchy of reference materials (tab. 2-7) (1, 4, 5). In this list the adjectives "primary" and "secondary" are used with a meaning which is different from that customarily used with standards (tab. 2-2).

Tab. 2-7. Hierarchy of reference materials.

Name of material	Synonym/Organization[1])
1. Primary Reference Material	
a) Pure Primary Reference Material	Standard Reference Material (SRM)/NBS
	Reference Material/ISO
	Primary Standard Material/ IFCC
b) Matrix Primary Reference Material	Standard Reference Material (SRM)/NBS
	Reference Material/ISO
2. Secondary Reference Material	Reference Material/ISO
	Control Material/IFCC
	Reference Material (Control)/NCCLS

[1]) Organizations:

IFCC	International Federation of Clinical Chemistry
ISO	International Organization for Standardization
NBS	National Bureau of Standards, USA
NCCLS	National Committee for Clinical Laboratory Standards, USA

Primary reference materials

This term is to be reserved for those reference materials which are prepared and whose properties are characterized with methods which insure that the best possible estimate of the "true value" is obtained. Thus a primary reference material would have to be analyzed with a definitive method (see Section 3), if available; otherwise other independent methods would have to be used for which extensive tests have shown that the results obtained are best possible estimates of the "true value". Primary reference materials in this sense are the standard reference materials of the NBS. However, control specimens with a biological matrix which have been prepared with great care and in which the analytes have been assayed with a definitive method would also be referred to as primary reference materials; examples of primary reference materials are given in Section 4.

Secondary reference materials

This term is used both for pure substances and for reference materials with all kinds of matrix (tab. 2-1).

For pure substances, purity is to be tested and confirmed by comparison with primary reference materials. For secondary reference materials with a more or less complex matrix, concentration is to be determined, if this is possible, using a reference method and primary reference materials or using methods which have been shown to be equally reliable. The most important use of such secondary reference materials is in calibration.

Control reference materials

This term refers to those materials which were described earlier as control specimens. They should have the properties already discussed. They are poorly suited or unsuitable for calibration of methods which are not completely specific.

3. Reference Methods

There are two different definitions of reference methods, each including a requirement of good precision (4, 24). The IFCC EPNPQC (2) defined a reference method as:

"A method which after exhaustive investigation has been shown to have negligible inaccuracy in comparison with its imprecision"

The NBS definition, on the other hand, contains additional requirements (1, 5):

"A reference method is an analytical method whose inaccuracy and imprecision are small enough as demonstrated by direct comparison with the definitive method, and whose low incidence of susceptibility to known interferences is (so) thoroughly documented that the stated end-purposes of the reference method may be achieved".

The additional requirements of the NBS are:

1. Testing of the reference method by comparison with a definitive method (2, 5).
2. Proof of low susceptibility of the method to interference factors.

   The NBS suggests that a definitive method be defined as "an analytical method that is capable of providing highest accuracy among all methods for determining that analyte, and its accuracy must be adequate for its stated end purposes" (5).

Uses and limitations

The purpose of a reference method (5, 25) is to provide, within specified confidence limits, good estimates of the "true value" of a constituent. These estimates are the basis of improvement of comparability of analytical results. Therefore reference methods should be used only in reference laboratories under competent directorship with well trained and highly motivated personnel. Reference methods must be transferable.

Reference methods are needed:

1. for determination of the best estimate of the "true value" of a constituent with specified confidence limits in standards and control specimens to be used in routine analysis.
2. as a means of monitoring the accuracy and specificity of routine methods and comparing routine methods.

Reference methods are not suitable for use in routine analysis because of the large specimen volume needed, the additional work involved and the highly qualified personnel and large amount of time needed for each analysis. A routine method and a reference method can be based on the same analytical principle, but with a reference method additional measures must be taken to attain the precision and accuracy required.

3.1 Reference methods in the strict sense

Reference methods which fit the definition of the NBS are referred to as "reference methods in the strict sense". These reference methods must be tested with a definitive method or with several equally good methods which are independent of one another.

In order to develop a candidate method into a reference method, the following must be available or agreement established in the area listed:

1. A primary standard with highest possible purity, e. g. a standard reference material of the NBS.
2. A method with good precision which is a suitable candidate reference method. Selection should be made with the following points in mind:
   a) The principle of measurement should have a solid theoretical basis.
   b) The following characteristics of the analytical system should be adequate for development into a reference method: precision and accuracy, linear range of measurement, specificity, sensitivity, stability.
3. Agreement on the acceptable deviation from the "true value".
4. Qualified reference laboratories for evaluation of transferability of the reference method.
5. A comprehensive system for quality control.
6. The possibility of consultation with a statistician during planning and evaluation of the investigations.
7. A group of experts to advise the principal investigator before starting and during the whole period of development and testing.
8. A definitive method or an equally good method whose accuracy is twice that required of the method being tested.

The development and testing of a reference method and its use are demonstrated using the example of calcium determination in serum. The principle of measurement underlying the definitive method needed for testing is compared with that of the reference method.

Reference method for determination of calcium in serum

Prior to the development of this reference method (25) all of the above criteria were fulfilled or agreement established:

1. The NBS standard reference material calcium carbonate (SRM 915) was available for use as the primary standard.
2. The atomic absorption method of *Pybus* et al. (27) met all of the above requirements and was selected as the candidate method.
3. A desirable maximum deviation of 0.5%, with a deviation of 1.0% probably attainable, was set as the accuracy target.
4. Eight qualified reference laboratories were available for the comparative investigations.
5. A complete quality control system was available.
6. The NBS provided statistical advice.
7. An expert advisory group was formed.
8. A definitive method for calcium determination was available at NBS. It is based on the principle of isotope dilution followed by mass spectrometry.

Each individual step of the reference method has been described in detail (25, 26). When reference analyses are made, strict adherence to these procedures is essential. No simplifications or changes of any other kind are permitted. Examples of the negative effects of such changes on precision and accuracy are given.

The following points in the protocol are of particular importance:

1. There are detailed specifications for the purity of both chemicals and water and for the preparation of reagents.
2. All volumetric glassware, i. e. the 10-ml volumetric pipette and the 500-ml volumetric flasks, should meet the highest specifications of the NBS, Class A. Cleaning and drying procedures are specified in detail.
3. Standard solutions and specimens are diluted 50-fold. The same 10-ml volumetric pipette (to deliver) is used each time to avoid errors due to differences in drainage times of serum and aqueous solutions and to eliminate errors due to volumetric differences between pipettes. Dilution takes place in a 500-ml volumetric flask.
4. Since among the eight reference laboratories originally evaluating the method three different models of double-beam atomic absorption spectrophotometer were used under different laboratory conditions, it was necessary for each laboratory to determine and then specify optimal conditions, including flame condition in particular.
5. Primary standard solutions are used for calibration, and the readings are recorded. The specimens are measured by bracketing, i. e. by taking a reading on the specimen and then on the two standard solutions with concentrations just above and below this reading, and then interpolating. A measurement is valid only if the difference between consecutive readings on the same standard is less than one percent (tab. 3-1).
6. Calculation of the results for the individual measurements is done by both mathematical and graphic interpolation.
7. The results of 10 valid measurements on a given specimen are compiled and the mean determined.

Definitive method for determination of calcium in serum

This method is based on the principle of isotope dilution with a stable calcium isotope followed by calcium isotope measurement utilizing mass spectrometry. The method has been described in detail elsewhere (25, 28). It consists of the following steps:

1. Spiking of the specimen with the stable isotope ^{44}Ca; chemical fusion of the specimen with HNO_3, and separation of interference factors which might affect mass spectrometry.
2. Careful measurement of the isotope ratio ($^{40}Ca/^{44}Ca$) in the serum sample and in the spiked serum sample using mass spectrometry.
3. Determination of the blank value resulting from impurities from the reagents and the fallout during chemical reactions and separation operations.
4. Determination of the new isotope ratio following spiking, calculation of the calcium concentration in the specimen.

3.2 Provisional reference methods

There are a good many constituents which are determined frequently in clinical chemistry but for which no definitive methods can be developed at present. This is because:

1. The analyte is not a defined, homogeneous substance (e. g. total protein in serum).
2. The analyte is a homogeneous substance, but there is no well-established analytical principle for its determination which could serve as the basis for the development of a definitive method (e. g. bilirubin in serum).
3. The analyte is not a substance but an activity (e. g. enzyme activities).

Tab. 3-1. Reference method for calcium determination: readings and valid measurements for standards and specimens.*)

Test No.	Specimen	Readings						
		Concentration of standard solutions of calcium (mmol/l)						
		2.00	*2.25*		*2.50*	*2.75*		*3.00*
1	Standards	1,084	1.222		1.355	1.486		1.608
2	Unknown specimen			0.001 (< 0.1%)	1.477		0.006 (0.4%)	
3	Standards				1.354	1.480		
4	Unknown specimen			0.003 (< 0.1%)	1.475		0.008 (0.5%)	
5	Standards				1.351	1.488		
6	Unknown specimen			0.031 (> 1.0%)	1.420			
				↖ Difference between readings of the same standard solution ↗				

*) Adapted from *Cali* et al. (26)

It is most expedient if provisional reference methods are developed and tested by experts at the request of a national or international scientific organization. The initiating body can then also work to achieve consensus agreement on such methods. The Center for Disease Control in Atlanta, Georgia, USA (CDC) has described a procedure for achieving such agreement that has proven useful in the USA (4). Examples of provisional reference methods are total protein determination in serum (29), including the required secondary standard (30), and the recommendations of the German Society for Clinical Chemistry (GSCC – Deutsche Gesellschaft für Klinische Chemie) for the standardization of the most frequent enzyme activity determinations (36, 37) in serum. In the Federal Republic of Germany about 95% of the enzyme activity determinations are currently carried out with standardized methods, which leads to good comparability of analytical results.

3.3 The hierarchy of methods

The methods discussed above can be ranked on the basis of their reliability, in particular their accuracy and specificity (tab. 3-2).

On this basis reference methods are followed by routine methods. Reference methods are clearly unsuited for routine analysis. But it is certainly possible to use reference methods or provisional reference methods to evaluate systematically the reliability criteria of established routine methods. The American Association for Clinical Chemistry (AACC) has developed a procedure for this (31). The methods tested according to this procedure are called selected methods (32). A similar procedure was developed by the Standards Commission of the GSCC. The final text was prepared jointly with the French Society for Clinical Chemistry and is now used by both societies. Methods tested in this way are called Selected Methods (Ausgewählte Methoden).

Because of the very small number of reference methods currently available, more complicated methods with adequate performance characteristics will have to be used instead of reference methods as the point of reference in evaluating candidates for Selected Methods.

Tab. 3-2. Hierarchy of analytical methods and analytical results.

Method	Result
–	"true value"
Definitive method	definitive value
Reference method	reference method value
Class A: tested with definitive method	
Class B: not tested with definitive method, but highly purified, defined standard available, reliability of method assured	
Class C: no homogeneous standards of known composition available, testing with definitive method not possible	
Routine method	assigned value
Class A: systematic error known (selected method = ausgewählte Methode)	
Class B: systematic error not known	

4. Experience with reference materials and reference methods

The improvement in the reliability and comparability of the results of routine clinical chemical analyses, which is possible if suitable reference materials and methods are used appropriately, is indicated below by a few examples. The uncertainty introduced into measurements if an unsuitable reference material is used is also discussed.

4.1 Reference materials

Both standards and control specimens are referred to as reference materials. But it is important that the reference materials used for calibration and those used for control of methods, which are not completely specific, be operationally separate; in addition they should have very different characteristics if they are to serve their different purposes optimally. For these reasons standards and control specimens are discussed separately below.

4.1.1 Standards

Primary standards

Primary standards (tab. 4-1) with specific characteristics (33) must be used, where this is feasible, by the reference laboratories in determining assigned values in control specimens for the Reference Commission of the GSCC. This common point of reference facilitates determination of the cause(s) of systematic errors in the results of a particular reference laboratory.

Secondary standards

In the total protein determinations in two liquid control specimens, it was found that when each reference labora-

Tab. 4-1. Primary standards for assigned value determination in control specimens (examples).

Constituent	Primary standard
Calcium	Fixanal® for atomic absorption (Riedel 38604)
Chloride	0.1 mol/l NaCl (Merck Titrisol 9945)
Iron	iron prepared by reduction p. a. (Merck 3819)
Glucose	*D*-glucose, anhydrous (Merck 8337)
Urea	urea p. a. (Merck 8487)
Creatinine	creatinine (Merck 5208)
Copper	copper(II)-oxide p. a. (Merck 2766)

tory used its own protein standard the participating laboratories had a wide range of means, i. e. between-laboratory precision was poor. When samples of the same albumin standard as described by *Peters* (29) were used, between-laboratory precision improved markedly (34) (tab. 4-2), while within-laboratory precision remained unchanged, with one exception. This shows how the introduction of a homogeneous, defined secondary standard can lead to improvement in the reliability, and thus the comparability, of analytical results.

4.1.2 Control specimens

For internal control of accuracy and for interlaboratory surveys, the Guidelines of the Medical Society of West Germany (Bundesärztekammer) specify that control specimens are to be used in which the concentration of the analytes is known. When routine methods which are more or less nonspecific are used to determine these concentrations, the values found differ from the "true value" (see tables 1-2 and 2-6) by the sum of the non-specific components and the effects of any procedural deficiencies. It is therefore necessary to determine assigned values for accuracy control specimens which are dependent on the method used.

Tab. 4-2. Improvement in between-laboratory precision by using the same secondary standard for determination of total protein in serum (Serum A and B: different standards; serum C and D: same standard).

Control specimen	Reference Laboratory						
Parameters	1	2	3	4	5	6	7
Serum A							
$\bar{x}$ (g/100 ml)	7.35	7.31	7.67	7.50	7.51	7.12	
CV (%)	1.4	1.65	4.65	0.9	2.3	2.1	
	$\bar{x}$ = 7.41 (g/100 ml)						
	$\overline{CV}$ = 2.4 (%)						
Serum B							
$\bar{x}$ (g/100 ml)	6.84	7.02	6.68	7.17	6.72	6.93	
CV (%)	0.95	2.25	3.95	2.2	2.0	1.35	
	$\bar{x}$ 6.92 (g/100 ml)						
	$\overline{CV}$ = 2.6 (%)						
Serum C							
$\bar{x}$ (g/100 ml)	6.48	6.59	6.73	6.55	6.73		
CV (%)	1.3	1.51	1.32	2.75	1.47		
	$\bar{x}$ = 6.62 (g/100 ml)						
	$\overline{CV}$ = 1.7 (%)						
Serum D							
$\bar{x}$ (g/100 ml)	5.28	5.23	5.21	4.99	5.18	5.24	5.18
CV (%)	1.5	3.1	1.9	3.4	2.9	1.9	1.2
	$\bar{x}$ = 5.17 (g/100 ml)						
	$\overline{CV}$ = 1.9 (%)						

Each laboratory made 30 determinations in different runs for each control specimen. The mean, standard deviation from day to day and coefficient of variation were calculated from the analytical results.

$\bar{x}$ mean of the means

$\overline{CV}$ mean of the coefficients of variation for a given control specimen

Assigned value determination in accuracy control specimens

On the basis of experimental studies and statistical evaluation of over 200 lots of control specimens (23, 25), the following points have been found to be of importance in the planning and evaluation of assigned value determinations (35):

1. In every control specimen lot the concentration of the analyte is unknown; furthermore, the composition and structure of the matrix vary from lot to lot.
2. Several methods with different specificity and different procedural deficiencies are usually available for the analysis of a particular constituent.
3. The conditions of measurement for a given method are by no means identical in different laboratories; at best they are equivalent.

Therefore the variability of the analytical results for a given constituent depends on:

1. the control specimen,
2. the method, and
3. the laboratories involved.

Furthermore, the analytical results from different laboratories can vary widely.

The result is that the analytical results which are obtained for samples of the same control specimen analyzed with the same method but in different laboratories cannot generally be regarded as a random sample of the same whole. It is thus not generally possible to assume a common probability distribution for such results. Rather, it is advisable to determine confidence limits without making any assumptions about the distribution of the analytical results and to call the location parameter of these confidence limits the assigned value.

Since both the volume of a homogeneous lot of a control specimen and the financial resources available for assigned value determination are limited, the experimental design used is a compromise which provides the maximum amount of information for the available funds. The following procedure has proven useful (fig. 4-3):

At least three laboratories participate in the assigned value determination for each constituent and each method. Each laboratory carries out duplicate determinations on 15 successive working days for lyophilized control specimens or 10 successive working days for liquid control specimens. The analytical results are

Sicherung der Digoxin-Therapie mit Merckotest®

Emit®-cad

Emit® = eingetragenes Warenzeichen der Firma Syva, Palo Alto, USA
cad = cardioactive drug

Die Blutspiegelkontrolle eines digitalisierten Patienten ist von immer größer werdender Bedeutung. Ihre Durchführung macht die Digoxin-Therapie sicherer.

Bei dem Merckotest® Emit®-cad Digoxin handelt es sich um einen homogenen, nicht radioaktiven Enzym-Immuno-Assay zur präzisen Bestimmung im Bereich von 0,5 bis 4,0 ng/ml.

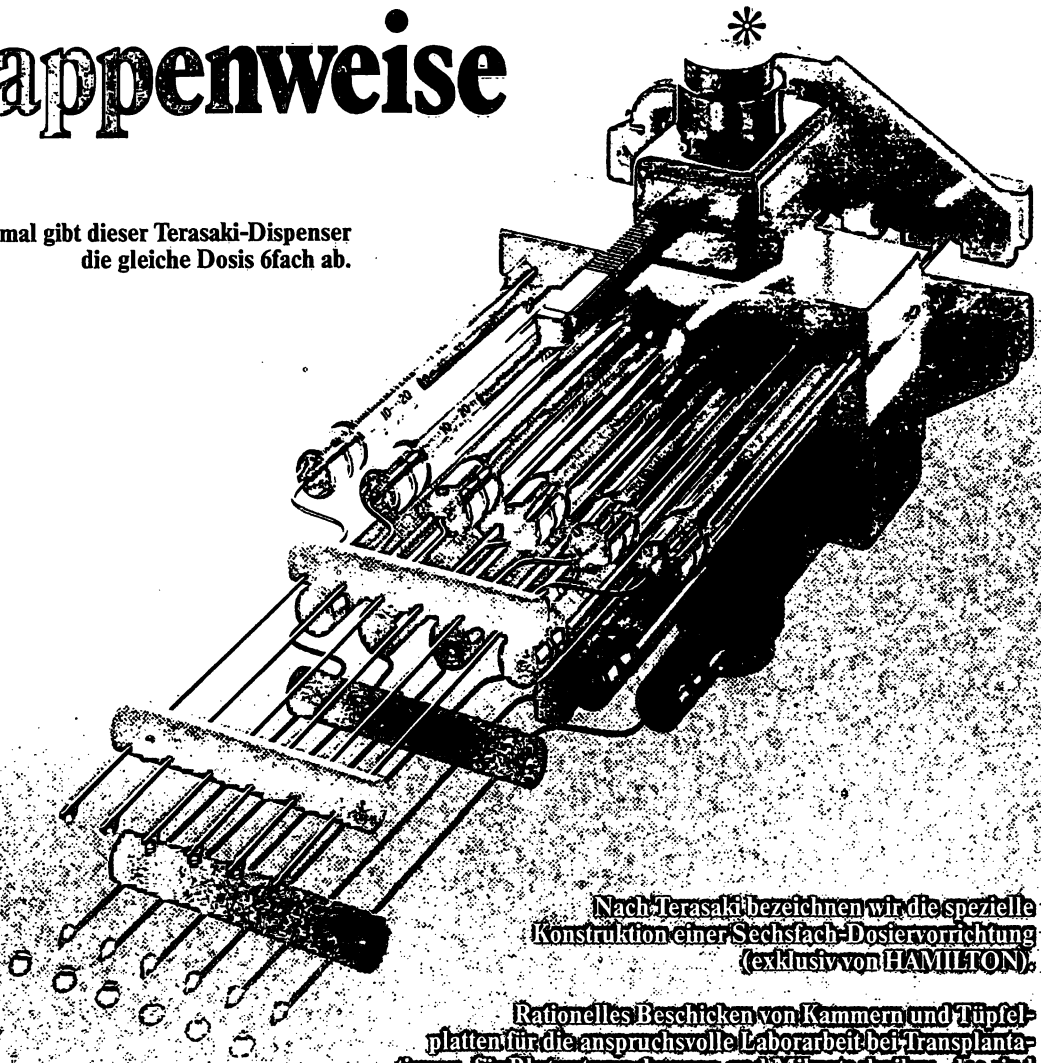
Bitte fordern Sie ausführliche Unterlagen an.

E. Merck · Frankfurter Straße 250 · D-6100 Darmstadt 1

Diagnostica MERCK

etappenweise

✳ 50mal gibt dieser Terasaki-Dispenser die gleiche Dosis 6fach ab.



Nach Terasaki bezeichnen wir die spezielle Konstruktion einer Sechsfach-Dosiervorrichtung (exklusiv von HAMILTON).

Rationelles Beschicken von Kammern und Tüpfelplatten für die anspruchsvolle Laborarbeit bei Transplantationen, für Blutuntersuchungen und Mikrotechniken, das sind die Haupt-Einsatzgebiete des Terasaki-Dispensers. Die Verbindung mit verschiedenen Spritzentypen von HAMILTON (von 50 bis 500 µl) ermöglicht eine Vielzahl von Volumenvariationen. Beim Terasaki betragen die Kanülenabstände einheitlich je 6,4 mm, wir können Ihnen jedoch auch Mehrfach-Dosiergeräte mit je 9 mm Kanülenabständen liefern.

Schreiben Sie bitte an Hamilton Bonaduz AG, CH-7402 Bonaduz, Schweiz, oder direkt an unsere Niederlassung in Deutschland.

HAMILTON

Hamilton Deutschland GmbH
D-6100 Darmstadt, Postfach 110427, Otto-Röhm-Strasse 74,
Telefon (06151) 8 50 85/86, Fernschreiber 419684

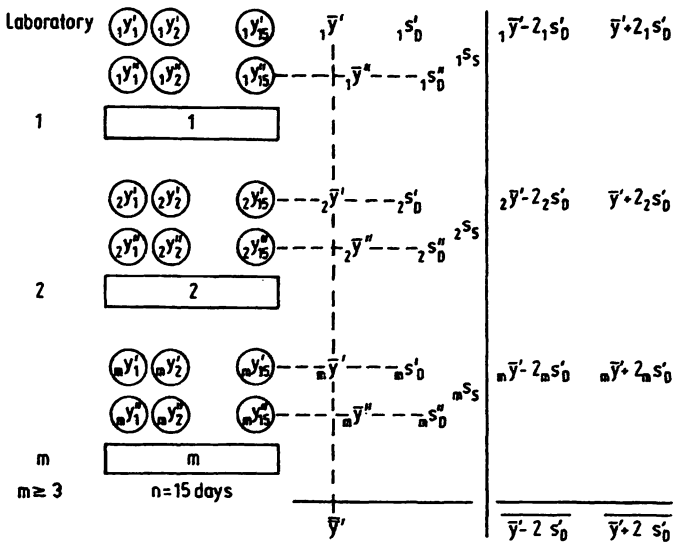


Fig. 4-3. Experimental design for assigned value determination.

recorded in a form suitable for transfer to punch cards. A computer is then used for statistical analysis.

For each reference laboratory the frequency distributions for the first and second results of a duplicate determination are determined separately and the means and standard deviations from day to day determined. Then the standard deviation of the series is determined from the differences in the pairs of values. This permits evaluation of the results of the individual laboratories (tab. 4-4).

The results of all those laboratories which have used the same method are combined into two frequency distributions – with separate distributions for the first and second results (tab. 4-5). The values for the parameters listed above are determined and in addition the median is found. Now suitable limits including 95% of the results are marked on the frequency distributions. The location parameter is also marked, the midpoint on the scale between the limits usually being used for reasons of symmetry. This location parameter is designated as the assigned value, and the limits specified above become the confidence limits.

If the total number of analyses is kept constant but the number of reference laboratories is increased and the number of duplicate determinations in the individual laboratories reduced accordingly, then the value of the location parameter remains virtually unchanged, whereas the confidence limits are often farther apart.

4.2 Reference methods

The reliability of assigned value determination in control specimens according to the procedure just described was tested effectively as follows: in samples of the same control specimen the concentration of several constituents is determined according to the experimental design for assigned value determination

(a) with a particularly reliable routine method, i. e. with a selected method, or with a candidate for a selected method,

(b) with a reference method and

(c) possibly also with a definitive method. The results of these three methods are then compared.

The possibility for such a comparison arose for a liquid bovine serum which was prepared at the request of the World Health Organization (WHO) by the Clinical Chemistry Division (Director: *David D. Bayse*) of the CDC. For the constituents listed (tab. 4-6), the definitive values were determined at the NBS, the reference method values at the CDC and the assigned values in the Department of Clinical Chemistry of the Max Planck Institute for Psychiatry in Munich, as Research and Reference Center of WHO, together with the Reference Commission of the GSCC.

The good agreement of the assigned values with the definitive values and the reference method values shows just how good agreement can be, and it documents the reliability of the procedure used for assigned value determination.

The comparability of the results of routine clinical chemical analyses done in physicians' offices can also be improved markedly if suitable standards and control specimens are used, as has been demonstrated in interlaboratory surveys (tab. 4-7).

Conclusions

1. Where possible, analytical methods in clinical chemistry should be calibrated and tested with highly purified, defined standards.
2. Primary standards should be used if at all possible. If none are available, then secondary standards with a detailed description of preparation and determination of concentration are recommended.
3. In addition, for the ongoing monitoring of performance, control specimens are needed whose matrix is so similar to the matrix of the specimens analyzed, e. g. patient specimens, that any procedural deficiencies or interference factors due to this matrix will be detected early.
4. Reference methods are needed both to make a reliable estimate of the "true value" of secondary standards and accuracy control specimens and in comparing routine methods.
5. Ideally these reference methods are tested with definitive methods. If a definitive method cannot be developed or is not available, provisional reference methods should be developed and used.

Tab. 4-4. Computer printout showing the analytical results for each reference laboratory and related statistics.

DEUTSCHE GESELLSCHAFT FÜR KLINISCHE CHEMIE ** ZENTRALE REFERENZINSTITUTION ** ABT. FUER SOLLWERTERMITTLUNG

SOLLWERTERMITTLUNG IN DER RICHTIGKEITSKONTROLLPROBE

REGISTRIERNUMMER : 10100
WARENZEICHENNAME :
HERSTELLER :
CHARGE : 433
AUSDRUCKDATUM : 15. 11. 74

BESTANDTEIL 17 PROTEIN EINHEIT G/100ML

METHODE BIURET-REAKTION MIT PROBENLEERWERT

2.		6.		11.		50.		51.	
1. ERGEBNIS	2. ERGEBNIS	1. ERGEBNIS	2. ERGEBNIS	1. ERGEBNIS	2. ERGEBNIS	1. ERGEBNIS	2. ERGEBNIS	1. ERGEBNIS	2. ERGEBNIS
5.5 0	5.5 0	5.5 0	5.5 1	5.5 0	5.5 0	5.5 0	5.5 0	5.5 0	5.5 0
5.6 0	5.6 0	5.6 0	5.6 1	5.6 0	5.6 0	5.6 0	5.6 0	5.6 0	5.6 0
5.7 0	5.7 0	5.7 4	5.7 3	5.7 0	5.7 0	5.7 0	5.7 0	5.7 0	5.7 0
5.8 0	5.8 0	5.8 4	5.8 3	5.8 0	5.8 0	5.8 0	5.8 0	5.8 0	5.8 0
5.9 2	5.9 3	5.9 4	5.9 4	5.9 0	5.9 0	5.9 2	5.9 0	5.9 0	5.9 0
6.0 4	6.0 5	6.0 0	6.0 0	6.0 0	6.0 0	6.0 2	6.0 1	6.0 0	6.0 0
6.1 5	6.1 5	6.1 2	6.1 2	6.1 0	6.1 0	6.1 6	6.1 5	6.1 0	6.1 0
6.2 4	6.2 2	6.2 1	6.2 1	6.2 2	6.2 2	6.2 4	6.2 5	6.2 0	6.2 0
6.3 0	6.3 0	6.3 0	6.3 0	6.3 2	6.3 3	6.3 1	6.3 4	6.3 2	6.3 2
6.4 0	6.4 0	6.4 0	6.4 0	6.4 9	6.4 8	6.4 0	6.4 0	6.4 2	6.4 1
6.5 0	6.5 0	6.5 0	6.5 0	6.5 2	6.5 2	6.5 0	6.5 0	6.5 3	6.5 3
6.6 0	6.6 0	6.6 0	6.6 0	6.6 0	6.6 0	6.6 0	6.6 0	6.6 3	6.6 4
6.7 0	6.7 0	6.7 0	6.7 0	6.7 0	6.7 0	6.7 0	6.7 0	6.7 4	6.7 3
6.8 0	6.8 0	6.8 0	6.8 0	6.8 0	6.8 0	6.8 0	6.8 0	6.8 1	6.8 2
$\bar{x}$ = 6.0640	6.0573	$\bar{x}$ = 5.8667	5.8400	$\bar{x}$ = 6.3753	6.3760	$\bar{x}$ = 6.0933	6.1753	$\bar{x}$ = 6.5533	6.5733
S = 0.0827	0.0892	S = 0.1589	0.1920	S = 0.0769	0.0944	S = 0.1081	0.0940	S = 0.1552	0.1580
V = 1.3639	1.4733	V = 2.7079	3.2874	V = 1.2061	1.4806	V = 1.7748	1.5226	V = 2.3687	2.4031
$\bar{x}$ – 2S = 5.8986	5.8789	$\bar{x}$ – 2S = 5.5489	5.4560	$\bar{x}$ – 2S = 6.2215	6.1872	$\bar{x}$ – 2S = 5.8770	5.9873	$\bar{x}$ – 2S = 6.2429	6.2574
$\bar{x}$ + 2S = 6.2294	6.2358	$\bar{x}$ + 2S = 6.1844	6.2240	$\bar{x}$ + 2S = 6.5291	6.5648	$\bar{x}$ + 2S = 6.3096	6.3634	$\bar{x}$ + 2S = 6.8638	6.8893
S* = 0.0179		S* = 0.1653		S* = 0.0256		S* = 0.0930		S* = 0.0408	

ANZAHL DER DATEN: 15

Tab. 4-5. Computer printout summarizing the results of the reference laboratories.

DEUTSCHE GESELLSCHAFT FÜR KLINISCHE CHEMIE ** ZENTRALE REFERENZINSTITUTION ** ABT. FUER SOLLWERTERMITTLUNG

SOLLWERTERMITTLUNG IN DER RICHTIGKEITSKONTROLLPROBE

REGISTRIERNUMMER : 10100
WARENZEICHENNAME :
HERSTELLER :
CHARGE : 433
AUSDRUCKDATUM : 15. 11. 74

BESTANDTEIL 17 PROTEIN EINHEIT G/100ML

METHODE BIURET-REAKTION MIT PROBENLEERWERT

ZUSAMMENFASSUNG

1. ERGEBNIS	2. ERGEBNIS	
5.5 0	5.5 1	
5.6 0	5.6 1	
5.7 4	5.7 3	
5.8 4	5.8 3	
5.9 8	5.9 7	$\leftarrow \bar{\bar{x}} - 2\bar{s}$
6.0 6	6.0 6	
6.1 13	6.1 12	
6.2 11	6.2 10	
6.3 5	6.3 9	
6.4 11	6.4 9	$\leftarrow \bar{\bar{x}} + 2\bar{s}$
6.5 5	6.5 5	
6.6 3	6.6 4	
6.7 4	6.7 3	
6.8 1	6.8 2	

Assigned value: 62 g/l

Confidence limits: 57 – 67 g/l

	G/100ML	G/L
$\bar{\bar{x}}$	= 6.1975	= 61.9746
$\bar{s}$	= 0.1209	= 1.2094
$\bar{\bar{x}} - 2\bar{s}$	= 5.9556	= 59.5557
$\bar{\bar{x}} + 2\bar{s}$	= 6.4393	= 64.3935
MEDIAN	= 6.2000	= 62.0000

ANZAHL DER TEILNEHMER: 5

Tab. 4-6. Comparison of definitive values, reference method values and assigned values in a liquid control specimen prepared by the CDC.

Constituent	Definitive Value	Reference Value	Assigned Value		Method
	NBS	CDC	GSCC (95% confidence limits)		
Calcium (mmol/l)	2.27	2.25	2.25	(2.05–2.45)	Atomic absorption
			2.125	(2.00–2.25)	Flame emission with acetylene
Chloride (mmol/l)	100.2	100.6	101	(98–104)	Coulometric
Glucose (mg/dl)	98.0	97.5	100	(94–106)	Hexokinase
Potassium (mmol/l)	4.79	4.82	4.85	(4.7–5.0)	Flame emission with Li-guideline
			4.75	(4.6–4.9)	without Li-guideline
Sodium (mmol/l)	142.4	142.4	143	(140–146)	with Li-guideline
			141	(138–144)	without Li-guideline
Magnesium (mmol/l)	0.87	0.87	0.9	(0.825–0.975)	Atomic absorption spectrophotometry

CDC Center for Disease Control
NBS National Bureau of Standards
GSCC German Society for Clinical Chemistry

Tab. 4-7. Comparability of the results of routine clinical chemical analyses from reference laboratories and from participants in interlaboratory surveys.
2nd interlaboratory survey of the Bavarian Panel Doctors Association, October 1975.

Between-Laboratory Precision (Coefficients of Variation)				Other Interlaboratory Surveys	
Constituent (Method)	Specimen	CV_R (%)	CV_P (%)	CV_{USA} (%)	CV_{UK} (%)
Glucose (GOD-Perid)	A	3.6	6.8	9.6	7.4
	B	3.6	6.7		
Urea (Urease/ *Berthelot* reaction)	A	4.6	10.9	9.4	10.6
	B	9.1	12.9		
Total protein (Biuret method)	A	3.5	5.9	–	5.1
	B	3.7	6.0		
Potassium (Eppendorf flame photometry)	A	3.3	4.3	3.0	3.2
	B	4.3	5.2		

CV_R Reference laboratories — CV_{USA} — USA (ACP)
CV_P Participants — CV_{UK} — England (Wellcome)

6. For methods which are not completely specific, the monitoring for systematic errors by comparison of the analytical results with the best estimate of the "true value" is frequently not effective enough. A comparison with an assigned value containing the same nonspecific components may be more effective.
7. Optimal standards, control specimens and reference methods are prerequisites for improvement in reliability and thus for the comparability of analytical results in clinical chemistry. Therefore special attention must be given to their further development.

References

1. Cali, J. P. (1975), International criteria for reference materials. In: Quality Control in Clinical Chemistry (Anido, G., van Kampen, E. J., Rosalki, S. B. & Rubin, M., eds.) Walter de Gruyter, Berlin, New York, pp. 369–383.
2. Büttner, J., Borth, R., Boutwell, J. H., Broughton, P. M. G. & Bowyer, R. C., Definitive recommendations on quality control in clinical chemistry, Part 1. General principles and terminology. Clin. Chim. Acta (in press), J. Clin. Chem. Clin. Biochem. (in preparation).
3. Büttner, J., Borth, R., Boutwell, J. H., Broughton, P. M. G. & Bowyer, R. C., Definitive recommendations on quality control in clinical chemistry, Part 2. Assessment of analytical methods for routine use. Clin. Chim. Acta (in press), J. Clin. Chem. Clin. Biochem. (in preparation).
4. Position Paper, Center for Disease Control. In: A National Understanding for the Development of Reference Materials and Methods for Clinical Chemistry. Proceedings of a Conference (Boutwell, J. H., ed.) American Association for Clinical Chemistry, Washington, D. C. 1978, pp. 163–239.
5. Position Paper, National Bureau of Standards. In: A National Understanding for the Development of Reference Materials and Methods for Clinical Chemistry. Proceedings of a Con-

ference (Boutwell, J. H., ed.) American Association for Clinical Chemistry, Washington, D. C. 1978, pp. 289–311.
6. Position Paper, American Association for Clinical Chemistry. In: A National Understanding for the Development of Reference Materials and Methods for Clinical Chemistry. Proceedings of a Conference (Boutwell, J. H., ed.) American Association for Clinical Chemistry, Washington, D. C. 1978, pp. 47–92.
7. Position Paper, American Society of Clinical Pathologists. In: A National Understanding for the Development of Reference Materials and Methods for Clinical Chemistry. Proceedings of a Conference (Boutwell, J. H., ed.) American Association for Clinical Chemistry, Washington, D. C. 1978, pp. 93–142.
8. Stamm, D. (1974), Calibration and quality control materials. Z. Klin. Chem. Klin. Biochem. *12*, 137–145.
9. Stamm, D. (1973), Spezifität von Analysenmethoden. In: Optimierung der Diagnostik (Lang, H., Rick, W. & Róka, L., eds.) Springer-Verlag, Berlin, Heidelberg, New York, pp. 187–201.
10. Pirke, K. M. & Stamm, D. (1972), Vergleich der Spezifität einfacher Methoden zur Bestimmung von Plasmacortisol, 1. Mitteilung. Porter-Silber-Methode und fluorometrische Verfahren. Z. Klin. Chem. Klin. Biochem. *10*, 243–253.
11. Pirke, K. M. & Stamm, D. (1972), Vergleich der Spezifität einfacher Methoden zur Bestimmung von Plasmacortisol, 2. Mitteilung. Die Proteinbindungsmethode. Z. Klin. Chem. Klin. Biochem. *10*, 254–259.
12. Kaiser, H. (1965), Zum Problem der Nachweisgrenze. Z. Anal. Chem. *209*, 1.
13. Büttner, H., Hansert, E. & Stamm, D. (1974), Statistical analysis, control and assessment of experimental results. In: Methods of Enzymatic Analysis, 2nd ed. (Bergmeyer, H. U., ed.) Academic Press, New York, London, Vol. 1, pp. 318–395.
14. Richtlinien der Bundesärztekammer zur Durchführung von Maßnahmen der statistischen Qualitätskontrolle und von Ringversuchen im Bereich der Heilkunde. Deutsches Ärzteblatt *68*, 2228–2231 (1971).
15. Ausführungsbestimmungen und Erläuterungen zu den Richtlinien der Bundesärztekammer zur Durchführung der statistischen Qualitätskontrolle und von Ringversuchen im Bereich der Heilkunde. Deutsches Ärzteblatt *71*, 961–965 (1974).
16. Stamm, D. (1972), Qualitätskontrolle klinisch-chemischer Analysen. Ein Lernprogramm in 5 Lektionen für medizinisch-technische Assistentinnen, Studenten und Ärzte. Programmiert von Zielinski, J., Tulodziecki, G. und Mitarbeitern. Georg-Thieme-Verlag, Stuttgart.
17. IFCC Newsletter No. 14, June 1976.
18. Radin, N. (1967), What is a standard? Clin. Chem. *13*, 55–76.
19. Young, D. S. & Mears, T. W. (1968), Measurement and standard reference material in clinical chemistry. Clin. Chem. *14*, 929–943.
20. Meinke, W. W. (1971), Standard reference materials for clinical measurements. Anal. Chem. *43*, 28A–47A.
21. Büttner, J., Borth, R., Boutwell, J. H., Broughton, P. M. G. & Bowyer, R. C., Provisional recommendation on quality control in clinical chemistry, Part 3. Calibration and control materials. Clin. Chim. Acta (in press), J. Clin. Chem. Clin. Biochem. (in preparation)
22. National Committee for Clinical Laboratory Standards: Calibration reference materials and control materials in clinical chemistry. Approved standard ASC-2, October 1974. 771 E. Lancaster Avenue, Villanova, PA 19085.
23. Stamm, D. (1975), The determination of assigned values for control specimens. In: Quality Control in Clinical Chemistry (Anido, G., van Kampen, E. J., Rosalki, S. B. & Rubin, M., eds.) Walter de Gruyter, Berlin, New York, pp. 117–130.
24. Cali, J. P. (1976), Rationale for reference methods in clinical chemistry. Pure Appl. Chem. *45*, 75.
25. Cali, J. P., Mandel, J., Moore, L. J. & Young, D. S. (1972), A reference method for the determination of calcium. National Bureau of Standards Special Publication 260–36, Washington, D. C.
26. Cali, J. P., Bowers, G. N., Jr. & Young, D. S. (1973), A referee method for the determination of total calcium in serum. Clin. Chem. *19*, 1208–1213.
27. Pybus, J., Feldman, F. J. & Bowers, G. N., Jr. (1970), Measurement of total calcium in serum by atomic absorption spectrophotometry, with use of a strontium internal reference. Clin. Chem. *16*, 998–1007.
28. Moore, L. J. & Machlan, L. A. (1972), High accuracy determination of calcium in blood serum by isotope dilution mass spectrometry. Anal. Chem. *44*, 2291–2296.
29. Peters, T., Jr. (1968), Proposals for standardization of total protein assays. Clin. Chem. *14*, 1147–1159.
30. National Committee for Clinical Laboratory Standards: Total serum protein (bovine serum albumin). Approved Standard ASC-1, 1973. 771 E. Lancaster Avenue, Villanova, PA 19085.
31. Guidelines for authors and evaluators of chapters for selected methods of clinical chemistry. July 1977.
32. Cooper, G. R. & Young, D. S. (1973), Selected methods. Clin. Chem. *19*, 1207.
33. Merck, E. (1971), Merck Standards. E. Merck, Darmstadt.
34. Müller, B. & Stamm, D. (1973), Experiences with a protein standard in clinical chemistry. In: Organisation des Laboratoires-Biologie prospective (Siest, G., ed.) L'Expansion Scientifique Française, Paris, p. 543.
35. Hansert, E. & Stamm, D., Assigned value determination in accuracy control specimens. (in preparation)
36. Recommendations of the German Society for Clinical Chemistry (1970), Z. Klin. Chem. Klin. Biochem. *8*, 658–660.
37. Recommendations of the German Society for Clinical Chemistry. Experimental evidence (1972), Z. Klin. Chem. Klin. Biochem. *10*, 281–291.

Prof. Dr. Dr. D. Stamm
Head, Department of Clinical Chemistry
Max Planck Institute for Psychiatry
Kraepelinstraße 10
D-8000 München 40